Issues in Identifying Easter Effects in Economic Time Series

Kellie Wills
U.S. Census Bureau

Disclaimer

This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.

Abstract

To decide whether the regARIMA model for an economic time series should include an Easter effect, seasonal adjusters at the U.S. Census Bureau use the AICC test in X-12-ARIMA (based on Akaike's Information Criterion corrected for sample size), along with subject matter information. However, the AICC test may be sensitive to model problems such as outliers whose t-statistics are below the threshold of the X-12-ARIMA automatic outlier detection algorithm. One would expect that series affected by Easter will show patterns of X-11 extreme values or regARIMA outliers in March and April, particularly in years when Easter falls in March, but such patterns have not been systematically investigated.

This paper shows that analysis of X-11 extreme values and regARIMA outliers can contribute to better Easter effect decisions. We will first examine patterns of these outlying values in synthetic series with known Easter effects. We will then consider two sets of Census series: retail sales series, for which there is a generally accepted economic rationale for an effect in the week before Easter; and manufacturing shipments series, for which Easter effects have not been investigated as thoroughly. We will show that X-11 extreme values in March and April are useful in identifying false Easter effect detections from the AICC test.

Keywords: holiday effects, seasonal adjustment

1. Introduction

The most common moving holiday effect found in U. S. economic flow series is the Easter effect. For many retail sales series, levels of sales are elevated in the period just before the Easter holiday (which varies between March 22 and April 25). The Census Bureau's X-12-ARIMA program estimates Easter effects by means of a regARIMA model – a time series model specifying a regression mean function that includes a holiday effect, with an ARIMA model for its autocovariance function.

The Easter regressor in X-12-ARIMA follows the simplest model of Bell and Hillmer (1983); it assumes that the level of activity changes *w* days before the holiday for a specified *w*, and remains at the new level through the day before the holiday. For a given effect window w, the Easter regressor E(w,t) is

$$E(w,t) = \frac{1}{w} \times n_{w,t} - \mu_{w,t}$$

where $n_{w,t}$ is the number of the w days before Easter falling in month t, and $\mu_{w,t}$ is the long-run calendar month (or quarter) mean of $n_{w,t}$ / *w* corresponding to the first 400-year period of the Gregorian calendar, 1583-1982. The $\mu_{w,t}$ capture the seasonal component, and thus their removal yields a regressor which does not estimate effects belonging to the seasonal component of the series. In this paper, we will refer to E(w,t) as easter[w].

X-12-ARIMA can test for the presence of an Easter effect using AICC (also called the F-corrected AIC), which is a version of Akaike's Information Criterion with a correction for sample size. Among competing models for a given time series, the model with the smallest AICC value is the model preferred by the criterion. For more information on AICC, see Hurvich and Tsai (1989). When performing the AICC test, X-12-ARIMA estimates the regARIMA model separately with each of the regressors easter[1], easter[8] and easter[15], and without an Easter regressor. It then selects the regressor yielding the smallest AICC, or the model without a regressor if its AICC is the smallest.

Findley and Soukup (2000) demonstrated moving holiday effect identification using AIC values and graphs of out-of-sample forecast errors. Findley, Wills and Monsell (2005) evaluated the performance of the X-12-ARIMA AICC test in identifying Easter effects in synthetic series. Results from that paper showed that additive outliers in March and April readily induce false detections. The present paper continues the study of Easter effect identification by investigating whether extreme values and outliers can aid in accurate identification.

A seasonal adjustment with X-12-ARIMA can incorporate two separate methods for identifying outlying values that may unduly influence the adjustment. First, the program offers automatic outlier detection for the regARIMA model. Second, the X-11 procedure identifies extreme values which it downweights or zero-weights in calculating the tables (Ladiray and Quenneville 2001). One would expect that series affected by Easter will show patterns of one or both of these types of outlying values in March and April. Since Easter usually occurs in April, years when Easter falls in March may be of particular interest. The comparison between outliers and extreme values is interesting because the outliers (like the AICC test) depend on the underlying ARIMA model, and thus may be influenced by model quality. The X-11 extreme value procedure, on the other hand, is nonparametric and is influenced by the regARIMA model only through the forecasts from the model.

We will first examine patterns of outlying values in synthetic series with known Easter effects. We will then consider Census retail sales series, for which there is a generally accepted economic rationale for an effect in the week before Easter, to examine the relationship between outlying values and AICC test results. Finally, we evaluate the evidence for Easter effects in manufacturing shipments series.

2. March/April Outlying Values in Synthetic Series

Findley, Wills and Monsell (2005) created seven sets of synthetic time series with 30 series each. One set had no Easter effect; to the other sets, we added Easter effects with windows of 1, 8 or 15 days and regression coefficients of 0.01 (small effect) or 0.03 (moderate effect). The synthetic data consist of 12 years of monthly observations ending April 2002, and thus include three years when Easter occurred in March (1991, 1997 and 2002). We performed default X-11 extreme value identification with sigmalim = (1.5,2.5) for these series, without adjusting for Easter. In a separate run, we set the automatic outlier identification procedure in X-12-ARIMA to identify additive outliers, using a low critical value (critical = 2.5) and an automatically identified ARIMA model with no Easter effect. We used the low critical value because Findley, Wills and Monsell (2005) showed that outliers with t-statistics this low can result in false Easter identifications; we were mindful of the potential for frequent Type I errors (many time points identified as outliers) in this multiple comparison situation.

Table 1 shows the number of March/April extreme values identified by a default X-11 seasonal adjustment run in years when Easter was in March. (Because this data span includes only three March Easters, there can be at most six extreme values.) Table 2 shows the number of March/April extreme values in all years. All series with a moderate (0.03) easter[1] or easter[8] effect had two or more March/April extreme values in March Easter years, but only two-thirds of series with a 0.03 Easter[15] effect did. This may result from the frequent extension of the easter[15] window into March in April Easter years. All series with a 0.03 Easter effect of any length had more than four March/April extreme values, considering all years.

Note that the cutoff of two or more extreme values in March Easter years would result in several false identifications in series with no Easter effect. The counts in series with the smaller 0.01 effect frequently fell below the cutoffs of two or more extreme values in March Easter years and more than four in all years.

	Number of series			
# of extreme values	**0**	**1**	**2**	**>2**
No Easter effect	7	20	3	0
Easter[1], 0.01	0	11	11	8
Easter[1], 0.03	0	0	0	30
Easter[8], 0.01	0	15	10	5
Easter[8], 0.03	0	0	5	25
Easter[15], 0.01	1	16	11	2
Easter[15], 0.03	0	10	6	14

Table 1. March/April extreme values in synthetic series (three March Easter years).

	Number of series			
# of extreme values	**2**	**3**	**4**	**>4**
No Easter effect	19	9	2	0
Easter[1], 0.01	5	12	12	1
Easter[1], 0.03	0	0	0	30
Easter[8], 0.01	3	15	11	1
Easter[8], 0.03	0	0	0	30
Easter[15], 0.01	2	11	16	1
Easter[15], 0.03	0	0	0	30

Table 2. March/April extreme values in synthetic series (all twelve years).

We present results for additive outliers in Tables 3 and 4. Overall, the series with Easter effects were less likely to show an elevated number of outliers than an elevated number of extreme values.

Number of series

# of outliers	**0**	**1**	**2**	**>2**
No Easter effect	23	7	0	0
Easter[1], 0.01	7	12	11	0
Easter[1], 0.03	0	2	6	22
Easter[8], 0.01	10	13	7	0
Easter[8], 0.03	0	7	9	14
Easter[15], 0.01	11	14	5	0
Easter[15], 0.03	5	9	7	9

Table 3. March/April additive outliers in synthetic series (three March Easter years).

Number of series

# of outliers	**0**	**1**	**2**	**>2**
No Easter effect	19	4	6	1
Easter[1], 0.01	6	10	9	5
Easter[1], 0.03	0	2	5	23
Easter[8], 0.01	10	7	9	4
Easter[8], 0.03	0	4	7	19
Easter[15], 0.01	9	6	9	6
Easter[15], 0.03	1	0	5	24

Table 4. March/April additive outliers in synthetic series (all twelve years).

These results suggest that attempts to identify Easter effects using only the evidence of March/April extreme values would miss many small Easter effects. (By comparison, the AICC test very rarely fails to identify small effects in these series, using all 12 years of data.) Identification attempts using only the evidence of March/April additive outliers would also miss some larger effects. However, the results confirm that adjusters should in general expect to see patterns of outlying values, particularly extreme values, in series with Easter effects. The next section examines several Census data series in which the AICC test identified Easter effects, but which had few March/April outlying values.

3. March/April Outlying Values in Retail Series

The Census Bureau currently adjusts some retail sales series for Easter. Since these are real (not synthetic) series, we cannot be certain whether or not an Easter effect exists. However, in certain retail sectors there is a clear economic rationale for an Easter effect – specifically, an increase in sales in the week before Easter. Findley and Soukup (2000) describe identification of Easter effects in these series.

We examined 29 retail series at the 3, 4 and 5-digit North American Industrial Classification System (NAICS) code levels, 15 of which the Census Bureau adjusts for Easter based on the AICC test and subject matter information. The coefficients of the Easter adjustments range from 0.011 to 0.094, with a median of 0.029 (close to the moderate size effect in our synthetic series). The 29 series began January 1992 and ended March 2005, a period in which there were 3 March Easters. X-12-ARIMA version 0.3 automatically selected the ARIMA model (assuming multiplicative adjustment) and ran the AICC test for Easter, including trading day in the regression model and using the default critical value for outlier detection. (Note that as part of the version 0.3 automatic selection procedure, the coefficients of Easter regressors selected by the AICC test are checked for significance at the 5% level; insignificant regressors are removed from the model.) AICC selected an Easter effect regressor for 18 series.

We compared the results of the AICC test to counts of outlying values identified when Easter was not included in the model. To identify additive outliers, we lowered the critical value to 2.5. We identified extreme values in a separate run using default X-11 seasonal adjustment and automatically identified ARIMA models.

Tables 5-8 show the number of series for which AICC did and did not select an Easter effect, crosstabbed with the number of March/April extreme values and additive outliers.

Number of series

# of extreme values	**0**	**1**	**2**	**3**	**4**
AICC: no Easter	5	4	1	1	0
AICC: Easter	2	3	7	2	4

Table 5. AICC decision vs. number of March/April extreme values in three March Easter years.

Number of series

# of extreme values	**2**	**3**	**4**	**>4**
AICC: no Easter	2	4	1	4
AICC: Easter	1	1	2	14

Table 6. AICC decision vs. number of March/April extreme values in all 13+ years.

Number of series

# of outliers	**0**	**1**	**2**	**3**	**>=4**
AICC: no Easter	6	5	0	0	0
AICC: Easter	4	3	8	2	1

Table 7. AICC decision vs. number of March/April additive outliers in three March Easter years.

Number of series

# *of outliers*	**0**	**1**	**2**	**3**	**4**
AICC: no Easter	2	3	1	3	2
AICC: Easter	1	0	2	5	10

Table 8. AICC decision vs. number of March/April additive outliers in all 13+ years.

In section 2 (Tables 1-2), we found that synthetic series with a moderate size (0.03) Easter effect generally had two or more March/April extreme values in March Easter years and more than four in all years. Like the synthetic series, the retail series discussed in this section include three years when Easter fell in March, so it makes sense to apply the same cutoff of two or more extreme values in March Easter years. In Table 5, gray shading highlights seven series for which the AICC decision was at odds with this cutoff. Five series had fewer than two extreme values in March Easter years, yet AICC preferred Easter. Two series had two or more extreme values, yet AICC preferred a model without Easter.

The retail series are slightly longer than the synthetic series (13 years 3 months vs. 12 years), so the cutoff of more than four extreme values in all years is less directly applicable; however, we used it as a starting point. Gray shading in Table 6 highlights eight series for which the AICC decision is at odds with this cutoff. Four series had four or fewer extreme values in all 13+ years, yet AICC preferred Easter. Four series had more than four extreme values, yet AICC preferred a model without Easter.

There is overlap between the highlighted series in Table 5 and those in Table 6. Six series total met one or both of the following conditions, yet AICC preferred Easter:

- Fewer than two extreme values in March Easter years;
- Four or fewer extreme values in all years.

Five of these six series also had fewer than two additive outliers in March Easter years. The Easter effect in two of the six series became insignificant when one March or April additive outlier was added to the model, suggesting that one unusual March or April value led the test to incorrectly identify an Easter effect. (The Census Bureau does not include an Easter regressor when seasonally adjusting one of these series, based on subject matter considerations.) For a third series, the model with an Easter effect had larger out-of-sample forecast errors in the last three years of the series. (See section 4.3 for a description of the out-of-sample forecast error diagnostic.) The remaining two series were unusual because their Easter regression coefficients were negative, whereas all other retail series with an Easter effect have a positive coefficient; the Census Bureau does not adjust these two series for Easter. Thus, there is evidence that the AICC test made a false Easter detection in five of the six series for which the number of extreme values fell below the cutoffs determined from the synthetic series results. These are the same five series which also had fewer than two additive outliers in March Easter years.

Four other series met one or both of the following conditions, yet AICC preferred a model without Easter:

- Two or more extreme values in March Easter years;
- More than four extreme values in all years.

The March/April extreme values in three of these series, 447 Gas Stations, 444 Building Materials and Garden and 45431 Fuel Dealers, may reflect high spring variability related to weather; 447 Gas Stations is also affected by the volatility of gasoline prices. Another series, 44831 Jewelry, is currently adjusted for Easter; although the Easter effect is not identified using an automatic model, it is significant in the improved model used by the Census Bureau.

This analysis is exploratory, as we have not examined the appropriateness of the AICC test decision in all 29 retail series. However, our results suggest that outlying values (either regARIMA outliers or X-11 extreme values) may be useful in identifying false positives from the AICC test. Attempts to use outlying values to identify effects missed by the AICC test require caution, since some series will show high weather-related variability in spring months.

4. Evidence for Easter Effects in Manufacturing Shipments Series

The Manufacturers' Shipments, Inventories and Orders (M3) survey measures monthly activity in the U.S. manufacturing sector. Users of the shipments data have suggested that some of these series may be affected by Easter. In response, the Manufacturing and Construction Division (MCD) at Census currently adjusts 34 of the 87 shipments series for Easter. (Not all 87 series are published.) MCD used X-12-ARIMA AICC test results to determine which series to adjust, with ARIMA models selected by MCD statisticians.

We performed an exploratory investigation of the shipments series, independent of MCD's results, to take the perspective of an analyst examining these series for the first time. Our first step was to determine which series the AICC test identified as having Easter effects, using automatic ARIMA modeling. We modeled data starting January 1997 (after the basis for classification changed from Standard Industrial Classification (SIC) to

NAICS) and ending September 2005. This time period includes three March Easters in 1997, 2002 and 2005. We ran the version 0.3 automatic ARIMA model selection procedure (assuming multiplicative adjustment) and the X-12-ARIMA AICC test for both trading day and Easter, using the default critical value for automatic outlier detection. For 24 of the series, this "automatic" model included an Easter regressor; the difference between this result and the 34 series adjusted by MCD reflects differences between the automatically selected ARIMA models and MCD's models. Eight effects were identified as easter[1], seven as easter[8] and nine as easter[15].

We evaluated whether any of the 24 Easter effects were artifacts of unidentified outliers or other modeling problems. We considered patterns of extreme values and additive outliers in March and April and examined the out-of-sample forecast error diagnostic to check the effect of Easter regressors on forecasting performance. Finally, we considered the differences in extreme values, additive outliers and out-of-sample forecast errors for different Easter effect windows.

4.1 Extreme Value/Additive Outlier Patterns

For each of the 24 series, we checked for possible improvements to the automatically selected regARIMA models, based on the Ljung-Box Q and residual spectrum diagnostics. We felt that regARIMA outlier results would be more informative if the outliers were identified from a good model (see comments on 4541 Electronic and Mail Order retail in section 3). For one series, U31AVS Iron/Steel Mills, we found that the Easter effect was not significant with our improved model.

Most of the 23 remaining series had negative Easter coefficients, indicating a reduction in activity in the period before Easter. The negative coefficients ranged from –0.017 to – 0.082. The series for which the Easter coefficients were positive were U11BVS Dairy Product Manufacturing and U34GVS Semiconductor and Related Device Manufacturing. While it is plausible that shipments of dairy products increase in the period before Easter, the result for semiconductors is more difficult to explain.

We examined the outlying values identified with the improved models, with no Easter effect included in the model. For extreme values, we used default X-11. We set a critical value of 2.5 for additive outliers (AO). Seven of the series had March/April additive outliers in March Easter years; thirteen had one or more total March/April additive outliers. Due to the small number of outliers found, we also considered the number of outliers found when we set the automatic outlier identification procedure to seek both additive and level shift (LS) outliers. The results appear in Tables 9-10.

Number of series

	0	**1**	**2**	**>2**
Extreme values	4	6	8	5
AO and LS outliers	12	6	4	1

Table 9. Outlying values in 23 manufacturing shipments series with AICC-identified Easter effects (three March Easter years).

Number of series

	<2	**2-4**	**>4**
Extreme values	2	11	10
AO and LS outliers	11	11	1

Table 10. Outlying values in 23 manufacturing shipments series with AICC-identified Easter effects (all 8+ years).

4.2 Impact of Outliers

4.2.1 Easter effects sensitive to outliers

Results from Findley, Wills and Monsell (2005) showed that additive outliers in March and April readily induce false AICC Easter detections in synthetic series. For five of the 23 shipments series with automatically identified Easter effects – nearly one-quarter of the series – the Easter effect was no longer significant when the regARIMA model included a March or April outlier with a t-statistic below X-12-ARIMA's default threshold for automatic identification. The t-statistics for the outliers ranged from 2.44 to 3.27 in absolute value.

Tables 11-12 reproduce Tables 9-10, excluding the five series in which the Easter effect was sensitive to outliers. Comparison with Tables 9-10 shows that all five outlier-sensitive series had fewer than two March/April extreme values in March Easter years.

Number of series

	0	**1**	**2**	**>2**
Extreme values	3	2	8	5
Additive outliers	10	3	4	1

Table 11. Outlying values in 18 shipments series with AICC-identified Easter effects which were not sensitive to outliers (three March Easter years).

Number of series			
	<2	2-4	>4
Extreme values	2	7	9
Additive outliers	9	8	1

Table 12. Outlying values in 18 shipments series with AICC-identified Easter effects which were not sensitive to outliers (all 8+ years).

4.2.2 Easter effect identification with outlier critical value 2.5

Sensitivity of an Easter effect to an outlier suggests that a single somewhat unusual March or April data month has caused the AICC test to falsely identify an Easter effect. Since the five "false positives" among the shipments series involved outliers with t-statistics below the default threshold for automatic identification, we decided to examine how lowering the critical value for outlier detection affected Easter identification. We ran the AICC test on the original set of 87 series with an outlier critical value of 2.5. We used the ARIMA models previously identified by the automatic selection procedure and included only significant Easter regressors. Under these conditions, AICC identified Easter effects in all series in which they were previously identified, as well as six additional series. These included two each of easter[1], easter[8] and easter[15]. The effects in three of these series, with improved models, were sensitive to March outliers that were not identified in the automatic run; for another series, the effect was sensitive to an automatically identified outlier when other outliers were omitted. In two other series, Easter effects which were previously borderline significant became significant with additional outliers. These results suggest that lowering the outlier critical value during detection did not prevent false Easter effect detections.

4.3 Forecasting Performance

A technique for comparing the forecasting performance of models with and without Easter is the out-of-sample forecast error plot (Findley and Soukup 2000, Findley 2005), available in X-12-Graph (Hood 2002). X-12-ARIMA's history spec is used to obtain differences of the accumulating sums of squared forecast errors between the competing models for forecast leads of interest (in this case, 1 and 12 months). If the direction of the accumulating differences is generally upward, then the forecast errors are predominantly larger for the first model (here, the model without an Easter regressor), and we prefer the second model that includes an Easter regressor.

Seven of the 18 shipments series from the last section had out-of-sample forecast error diagnostics that clearly indicated a preference for the Easter effect. An example, U25AVS Pesticide, Fertilizer and Other Agricultural Manufacturing, is shown in Figure 1. Large improvements in forecast error due to the easter[8] regressor appear as step increases between March and April in 2003 and 2004 (April Easters) and between February and March in 2005 (a March Easter). All seven series with better forecast errors from the model with Easter had two or more March/April extreme values in March Easter years. For the remaining 11 series, forecast error diagnostics favored neither model.

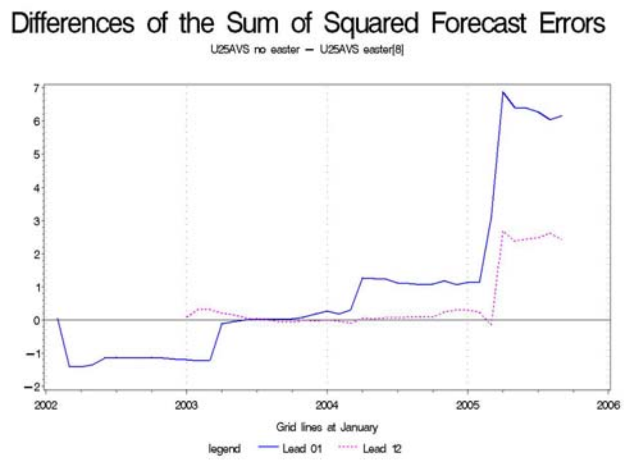

Figure 1. Out-of-sample forecast error diagnostic for U25AVS.

4.4 Impact of Easter Adjustment on Seasonal Adjustments

Figures 2 and 3 illustrate the impact of Easter adjustment on seasonally adjusted March and April data. Easter coefficients varied from 0.025 to 0.082 in absolute value, with a median of 0.037. Figure 2 shows the distribution of maximum and minimum Easter factors in the 18 series from section 4.2. For each of the 18 series, Figure 3 shows the maximum absolute percent difference between the seasonal adjustment including Easter adjustment and the seasonal adjustment without.

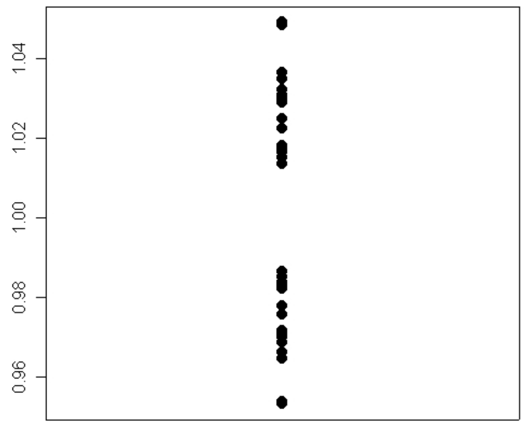

Figure 2. Ranges of maximum and minimum Easter factors in shipments series (18 series).

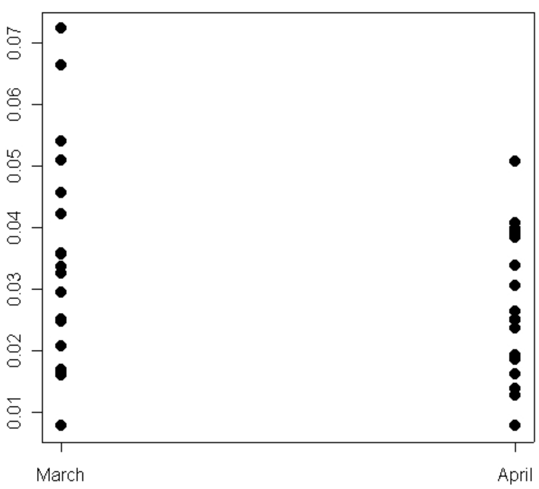

Figure 3. Maximum absolute percent difference between the seasonal adjustment including Easter adjustment and the seasonal adjustment without (18 series).

4.5 Length of Effect Window

For seven of the 18 series, the minimum AICC Easter window with the optimized model was easter[1], which suggests an Easter Saturday effect. However, given the Easter dates in the time span of the data, easter[1] is indistinguishable from easter[2] or easter[3], so that the regressor may actually be picking up a Good Friday effect. For six series, the minimum AICC Easter window was easter[8] and for the remaining five, it was easter[15].

We considered whether patterns of outlying values might provide information on what effect window is most appropriate. In previous sections, we considered outlying values identified when the model did not include Easter; in this section, we consider the outlying values identified when the model does include Easter. If a model with easter[8], say, results in more outlying values than a model with easter[1] for the same series, easter[8] may be modeling the effect less well. For each of the 18 series, we modeled the Easter effect as easter[1], easter[8] and easter[15] and examined the March and April extreme values identified with each model. For six series, the windows with the fewest extreme values differed from the minimum AICC window. We also examined March/April outliers (AO and LS, critical=2.5) identified with each model. For three series, the window with the fewest outliers differed from the minimum AICC window. (One of these series was also among the previous six.)

We examined the out-of-sample forecast errors for the eight series discussed in the previous paragraph. For five series, the out-of-sample forecast errors did not show a clear preference for any of the effect windows. For three other series, the minimum AICC window had the best out-of-sample forecast errors. Thus, for the shipments series, outlying values were not useful in identifying the most appropriate Easter effect window.

Conclusions

This paper has investigated whether analysis of X-11 extreme values and regARIMA additive outliers can contribute to better Easter effect decisions. Results from synthetic series with known Easter effects indicate that it would be impractical to identify Easter effects using only the evidence of March/April extreme values and/or additive outliers. However, the results confirm that adjusters should in general expect to see patterns of outlying values, particularly extreme values, in series with Easter effects. Results from Census Bureau retail series indicate that the X-12-ARIMA AICC test result broadly corresponds to the number of March/April extreme values and additive outliers: series with few extreme values/outliers tend not to have an Easter effect preferred by AICC, while series with many generally do. Analysis of the small number of series with an automatically identified Easter effect, but few outlying values, suggests that outlying values may be useful for detecting false Easter identifications from the AICC test. On the other hand, attempts to use outlying values to identify effects missed by the AICC test require caution, since some series will show high weather-related variability in spring months.

We applied these insights to the question of whether some Census Bureau manufacturing shipments series should be adjusted for Easter, as data users have suggested. With automatic modeling, the AICC test identified Easter effects in 28% of shipments series. However, nearly one-quarter of these effects became insignificant when the regARIMA model included a March or April outlier, suggesting a false detection. All of the series with outlier-sensitive effects had fewer than two March/April extreme values in March Easter years. Seven of the remaining series had out-of-sample forecast error diagnostics that clearly indicated a preference for the Easter effect; all had two or more March/April extreme values in March Easter years. Extreme values and additive outliers did not aid in identifying the most appropriate effect window for the series.

Our results indicate that X-11 extreme values in March and April – especially those in years when Easter falls in March – are useful in identifying false detections. In particular, a lack of these may be a sign that an outlier in March or April has induced a false Easter detection, an occurrence which Findley, Wills and Monsell (2005) and this paper show to be common. Lowering the critical value for outlier detection when performing the AICC test in manufacturing shipments series did not prevent these false detections. When using X-12-ARIMA, it is straightforward to check the extreme values identified by a default X-11 procedure, either in the output file or the diagnostic file, and our results suggest it is worthwhile to do so.

Acknowledgements

The author would like to acknowledge David Findley, Jonathan Holland, Kathy McDonald-Johnson and Michael Shimberg for insightful comments on the paper and presentation.

References

Bell, W.R. and Hillmer, S.C. (1983), "Modeling Time Series With Calendar Variation," *Journal of the American Statistical Association*, 78, 526-534.

Findley, D.F. (2005), "Asymptotic stationarity properties of out-of-sample forecast errors of misspecified regARIMA models and the optimality of GLS for one-step-ahead forecasting," *Statistica Sinica*, 15, 447-476.

Findley, D.F. and Soukup, R.J. (2000), "Modeling and Model Selection for Moving Holidays," *2000 Proceedings of the American Statistical Association,* Business and Economic Statistics Section, http://www.census.gov/ts/papers/asa00_eas.pdf.

Findley, D.F., Wills, K.C., and Monsell, B.C. (2005), "Issues in Estimating Easter Regressors Using RegARIMA Models with X-12-ARIMA," *2005 Proceedings of the American Statistical Association,* Business and Economic Statistics Section, http://www.census.gov/ts/papers/jsm2005bcm.pdf.

Hood, C.C. (2002), "X-12-Graph: A SAS/GRAPH Program for X-12-ARIMA Output, User's Guide for X-12-Graph Interactive for PC/Windows, Version 1.3," U. S. Census Bureau, U. S. Department of Commerce.

Hurvich, C.M. and Tsai, C. (1989), "Regression and Time Series Model Selection in Small Samples," *Biometrika*, 76, 297-307.

Ladiray, D. and and Quenneville, B. (2001), *Seasonal Adjustment with the X-11 Method.* New York: Springer-Verlag.